# Analysis of time-correlated single photon counting data

# Analysis of TCSPC data: a comparative evaluation of deterministic and probabilistic approaches

**Darren A Smith[1,2], Grant McKenzie[1], Anita C Jones[1], and Trevor A Smith[3]**

[1] EaStCHEM School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK
[2] Current address: School of Chemistry, University of Birmingham, Edgbaston, Birmingham, West Midlands, B15 2TT, UK
[3] School of Chemistry, University of Melbourne, Parkville, Victoria, 3010, Australia

E-mail: d.a.smith.3@bham.ac.uk

**Abstract.** We review various methods for analysing time-resolved fluorescence data acquired using the time-correlated single photon counting (TCSPC) method in an attempt to evaluate their benefits and limitations. We have applied these methods to both experimental and simulated data. The relative merits of using deterministic approaches, such as the commonly used iterative reconvolution method, and probabilistic approaches, such as the smoothed exponential series method (SESM), the maximum entropy method (MEM) and recently proposed basis pursuit denoising (BPDN, compressed sensing) method are outlined. In particular, we show the value of using multiple methods to arrive at the most appropriate choice of model. We show that the use of probabilistic analysis methods can indicate whether a discrete component or distribution analysis provides the better representation of the data.

## 1. Introduction

Time-correlated single photon counting (TCSPC) is a powerful method for the acquisition of time-resolved fluorescence data. It affords exquisite signal-to-noise over many orders of magnitude in intensity and time (from tens of picoseconds to microseconds). Commercial data analysis programs provide the ability to analyse data in terms of functions consisting of sums of a limited number of discrete exponential decay components, while more advanced programs also provide lifetime distribution analyses in one form or another. However, determining the most appropriate kinetic scheme with which to interpret the data is non-trivial.

The purpose of this paper is to evaluate the use of probabilistic analysis methods to gain an unbiased indication of the most appropriate model describing time-resolved fluorescence decay data. The intention is to provide an objective assessment of the benefits of using such an approach to determine a suitable model for the system under study. The results provide useful insight into the strengths and weaknesses of the various approaches investigated. Although the work in this paper is based on the time-correlated single photon counting method, the outcomes are relevant to all time-resolved fluorescence techniques.

The paper should also serve as a useful guide in the analysis of time-resolved data more generally. Fundamental analysis principles are outlined and practically demonstrated through the results that are presented. This should be of significant value to the growing time-resolved spectroscopy community, which is constantly striving to probe ever more complicated systems. Analysis of such systems can be extremely challenging, which can lead to misinterpretation of the data collected. It is hoped that the strategies outlined below will help to avoid such a fate.

### *1.1. Convolution and an ill-posed problem*

It is well known that the instrument response function (IRF) limits the achievable temporal resolution of TCSPC, and other time-resolved experiments. The IRF is borne of the finite width of the excitation pulse, as well as the temporal broadening caused by each of the components of the detection system (photo-detector, monochromator, and electronics), and the duration and substructure of the IRF results in a perturbation to the measured decay that cannot be neglected, particularly on short time-scales. The measured decay intensity, $D(t)$, is then a convolution of the IRF, $R(t)$, and the true fluorescence decay, $F(t)$;

$$D(t) = \int_0^t F(t')R(t-t')dt'. \qquad (1)$$

The convoluted nature of this relationship means the influence of the IRF cannot be removed by simply subtracting the IRF from the measured decay. The distortion of the true fluorescence decay by the IRF causes a serious problem for the analysis of TCSPC data; obtaining the true form of $F(t)$ requires inversion of equation (1), which is mathematically challenging because the problem is ill-posed [1-3]. Simply stated, this means that many solutions exist that can adequately describe the observed decay behaviour. Under these conditions the presence of noise can have significant influence on the solution obtained. Unfortunately, direct deconvolution strategies are generally found to be inappropriate for time-resolved fluorescence measurements [4]. For instance, Fourier transform approaches suffer from the discontinuity caused by truncation of the decay at long times (and also by the almost instantaneous excitation rise) in addition to the presence of noise [5, 6]. Simulated-annealing has been shown to provide promising deconvolution results [6]; however, this method is inherently slow due to the random nature of the optimisation.

Without a direct method to extract the true fluorescence decay from the measured decay data it has been necessary to develop strategies that can find reliable solutions to equation (1).

We should also point out that whilst the vast majority of fluorescence decay fitting in commercial programs and the literature is performed in terms of fluorescence *lifetimes*, it is often more physically meaningful to deal with *rate constants*. Any analysis method should be able to deal with both representations. For this reason, we discuss the scenarios in terms of both lifetimes and rate constants in this paper.

### *1.2. Iterative reconvolution (ItRe) of an assumed function*

Iterative (re)convolution is the most commonly used analysis method for TCSPC measurements because it is a robust technique that can quickly obtain solutions and, historically, it has been seen to be the most effective way to obtain reliable and accurate results [4, 5]. There are a number of commercial and open-source software packages that use iterative reconvolution, including FAST (Edinburgh Instruments Ltd.), FluorFit (PicoQuant), and Decay Fit (FluorTools, www.fluortools.com).

During fitting, the true fluorescence decay (which has been normalised here) is estimated by a lifetime (or rate constant) distribution;

$$\frac{F(t)}{F_0} = F_N(t) = \int_0^\infty p(\tau)e^{-\frac{t}{\tau}}\,dt, \qquad (2a)$$

$$F_N(t) = \int_0^\infty p(k)e^{-kt}\,dt, \qquad (2b)$$

where $\tau$ is the lifetime ($k$ is the rate constant) and $p(\tau)$ ($p(k)$) is the corresponding probability amplitude. For multi-exponential analysis – which is almost invariably used – $p(\tau)$ ($p(k)$) is represented by a weighted sum of $n$ discrete delta functions positioned at lifetimes $\tau_j$ (or rate constants $k_j$);

$$p(\tau) = \sum_{j=1}^{n} a_j\delta(\tau - \tau_j), \qquad (3a)$$

$$p(k) = \sum_{j=1}^{n} a_j\delta(k - k_j), \qquad (3b)$$

which gives,

$$F_N(t) = \sum_{j=1}^{n} a_j e^{-\frac{t}{\tau_j}}, \qquad (4a)$$

$$F_N(t) = \sum_{j=1}^{n} a_j e^{-k_j t}, \qquad (4b)$$

where $a_j$ is the contribution of the $j^{\text{th}}$ lifetime (rate constant) under the normalisation condition;

$$\sum_{j=1}^{n} a_j = 1. \qquad (5)$$

The resulting assumed decay function must be convolved with a measured (or estimated) IRF to allow comparison with the experimentally measured decay. Fitted parameters, $a_j$ and $\tau_j$ (or $k_j$), are iteratively improved until an adequate representation of the data is achieved.

As a point of clarification, it should be noted that all of the methods described in this work rely on iterative improvement of the fit of a convolved function; however, the term iterative reconvolution will be reserved for the deterministic type of analysis described in this section (*i.e.* the case where a *pre-determined* function is fitted to the experimental decay by the process of iterative reconvolution).

A significant limitation of the iterative reconvolution method is that it requires the assumption of a model before fitting (which is then validated afterwards). Ideally the system under study would comply with an intuitively obvious choice of model but, unfortunately, there is often little or no knowledge that can aid in proposing the definitively appropriate model. Without prior knowledge, the model is usually accepted or rejected based on its statistical validity after fitting using the null hypothesis [5]. Only if the evaluated model meets predetermined

criteria (*e.g.* a goodness-of-fit measure reaching a threshold value; see below) can it be accepted as a credible description of the system. It is important to realize that this does not necessarily mean that it is the correct model for the system; there may be a number of different models that satisfy the chosen criteria. In practice, the simplest possible model (*i.e.* a single-exponential decay function) is usually tested, before progressively more complexity is added until an acceptable fit has been achieved. Once an adequate fit has been established, there is generally no justification for assuming that a more complicated model would be beneficial or, indeed, appropriate. Some of the methods that can be used to determine the adequacy of a fit are given below.

*1.2.1. Evaluating quality of fit.* The goodness-of-fit is generally determined by least-squares analysis using the chi-squared statistic, $\chi^2$;

$$\chi^2 = \sum_{i=1}^{N_c} \left[ \frac{O(i) - E(i)}{\sigma(i)} \right]^2 = \sum_{i=1}^{N_c} \left[ \frac{D_{obs}(i) - D_{calc}(i)}{\sqrt{D_{obs}(i)}} \right]^2, \quad (6)$$

where $i$ represents the channel number; $N_c$ is the total number of channels in the fit; $O$ is the observed data (the measured decay, $D_{obs}$); $E$ is the estimate of the data based on the chosen decay function (the calculated decay after convolution with the IRF, $D_{calc}$), and $\sigma$ is the standard (or expected) deviation of the data. In Poisson statistics, which is typically assumed for fluorescence decay data collected by TCSPC, the variance of the data, $\sigma^2$, is equal to $D_{obs}$. The value of $\sigma^{-2}$ is known as the data weight, and the residuals of the fit, $r$, are properly weighted by $\sigma^{-1}$, that is;

$$r(i) = \frac{D_{obs}(i) - D_{calc}(i)}{\sqrt{D_{obs}(i)}}. \quad (7)$$

To account for the number of fitting parameters in the model, $n_p$, the $\chi^2$ value is typically normalised by the degrees of freedom in the fit, $\nu = N_c - n_p - 1$, to produce the reduced chi-squared statistic, $\chi_R^2$;

$$\chi_R^2 = \frac{\chi^2}{\nu} = \frac{\chi^2}{N_c - n_p - 1}. \quad (8)$$

Roughly speaking the $\chi_R^2$ value obtained for a fit can be interpreted in the following way: $\chi_R^2 \gg 1$ describes a poor fit; $\chi_R^2 > 1$ suggests that the model does not fully account for the observed behaviour; $\chi_R^2 \cong 1$ indicates that, within the expected variance, the observed fluorescence decay behaviour has been adequately matched by the model; finally, $\chi_R^2 < 1$ means that the data has been over-fitted, which can occur when there are too many variable parameters in the model, the error variance has been overestimated, or the fitting range is inappropriate. It should be noted that determining whether or not $\chi_R^2$ is close enough to unity is somewhat subjective and is dependent on the quality of data collected (which is influenced by IRF structure, photon yield, background level *etc.*). For most systems, $\chi_R^2 < 1.2$ is typically a reasonable indicator of an acceptable fit, but this should be checked against other criteria.

One of the most important (and challenging) aspects of the analysis of time-resolved data is the assessment of the residuals of the fit. While it is a helpful indicator of the overall quality of the fit, the chi-squared value only informs the experimentalist about the magnitude of the residuals and tells nothing about their distribution. The presence of structure within the residuals (*i.e.* a non-random distribution) represents an inability of the fitted model to accurately describe the data.

Birch and Imhof have provided a useful summary of the benefits of visually assessing properly weighted residuals [7]: it is possible to see *where* a fitted function does not match the data; the standardised weighting allows direct comparison between data sets of varying precision (signal-to-noise ratio); residuals are weighted by the standard deviation of the associated data and thus in a statistically significant way; finally, there is an intuitive relationship between the residuals and the associated chi-squared value.

Structure within the residuals can be subtle and difficult to spot (especially on short timescales). To aid in the assessment of the randomness of residuals it is possible to calculate their autocorrelation function (ACF) [8];

$$\mathrm{ACF}(i) = \sum_{m=1}^{N_c - i} \frac{(r_i - \langle r \rangle)(r_{m+i} - \langle r \rangle)}{\sum_{m=1}^{N_c} (r_i - \langle r \rangle)^2}, \quad (9)$$

where $\langle r \rangle$ is the mean residual value. The ACF can accentuate structure that is obscured within the weighted residuals. For instance, if there is underlying oscillatory behaviour within the residuals (perhaps due to electrical interference in the detector system) then this will be evident within the ACF.

Another way to investigate structure within the residuals is to calculate the Durbin-Watson parameter, $DW$;

$$DW = \frac{\sum_{i=2}^{N_c} (r_i - r_{i-1})^2}{\sum_{i=1}^{N_c} r_i^2}. \quad (10)$$

The Durbin–Watson parameter can range between 0 and 4 and is compared to lower ($DW_{L,\alpha}$) and upper ($DW_{U,\alpha}$) critical values, which depend on the significance level, $\alpha$, of the test. Deviation from the ideal value of 2 indicates the presence of structure within the residuals. For instance, the $DW$ parameter rejects the null hypothesis that residuals are **not** autocorrelated at the

95% significance level if $DW < 1.75$ (for a two-exponential decay) [9, 10]. Note that an "acceptable" $DW$ parameter ($DW > 1.75$, in this case) does not guarantee the residuals are uncorrelated (it only means that the Durbin–Watson test does not provide statistical evidence that they are autocorrelated ($DW < DW_{L,\alpha}$) or anticorrelated ($DW > DW_{U,\alpha}$) at the given significance level).

It is worth pointing out that the autocorrelation function and Durbin–Watson parameter are based on the same information as the weighted residuals analysis [7]. In other words, they present the same information in different ways, which often means that they are redundant in practical situations.

The Runs Test can be used to investigate whether or not the residuals are mutually independent (*i.e.* randomly distributed). It complements chi-squared analysis because it considers the sign of the error rather than the scale of the error. Here, a run is defined as a string of consecutive positive or negative residual values, with the next run beginning at a residual value of opposite sign. The test statistic, $Z$, is calculated as [11]

$$Z = \frac{R - R_e}{\sigma_R}, \tag{11}$$

where $R$ is the observed number of runs, and $R_e$ is the expected number of runs, calculated as

$$R_e = \frac{2n_1n_2}{n_1 + n_2} + 1, \tag{12}$$

where $n_1$ and $n_2$ are the number of positive and negative residual values and $\sigma_R$ is the standard deviation of the number of runs, calculated *via*

$$\sigma_R^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}. \tag{13}$$

For a random sequence of residuals $R$ should equal $R_e$ and so, in this ideal case, $Z$ diminishes to 0. The randomness of the residuals is rejected at the 95% significance level if $|Z| > 1.96$ [10]. Note that, similar to the Durbin-Watson test, $|Z| < 1.96$ does not guarantee the residuals are random (it only means the Runs Test cannot reject randomness at the given significance level, in this case $\alpha = 0.05$).

*1.2.2. Global analysis.* The reliability and robustness of a fit can be improved with the use of global analysis [1, 12, 13]. This approach requires simultaneous fitting of multiple decays that differ by some known parameter; typically the excitation or, more commonly, the emission wavelength is altered. The fluorescence decay is then described by:

$$F_N(t,\lambda) = \sum_{j=1}^{n} a_j(\lambda) e^{-\frac{t}{\tau_j(\lambda)}} \cong \sum_{j=1}^{n} a_j(\lambda) e^{-\frac{t}{\tau_j}}, \tag{14a}$$

$$F_N(t,\lambda) = \sum_{j=1}^{n} a_j(\lambda) e^{-k_j(\lambda)t} \cong \sum_{j=1}^{n} a_j(\lambda) e^{-k_j t}, \tag{14b}$$

where $a_j(\lambda)$ and $\tau_j(\lambda)$ ($k_j(\lambda)$) are wavelength-dependent amplitudes and lifetimes (rate constants). Generally the lifetimes (rate constants) are assumed to be wavelength-independent (since, in condensed phase, emission from the same excited state is observed, regardless of excitation or emission wavelength) and so can be constrained to have common (global) values when analysing multiple decays. This reduces the overall number of adjustable parameters in the analysis and militates against the problem of correlation between amplitudes and lifetimes.

Despite improving confidence in the final result of fitting, global analysis does not solve the fundamental problem of having to choose a model before fitting. For this reason global analysis will not be discussed further in this work, but for a comprehensive review and further reading see van Stokkum *et al.* and references therein [14].

*1.3. Probabilistic analysis*

From the above assessment it should be evident that iterative reconvolution cannot definitively provide the correct model for the system under study. Additionally, if a number of plausible solutions exist, it is not always clear which model is physically most appropriate. Therefore, to gain insight into the underlying dynamics of the decay process, it may be useful to use a probabilistic (or distributive) approach that makes no initial assumptions about the physical model of the system. In this case $p(\tau)$ ($p(k)$) in equation (2) is not limited to any particular functional form. For this strategy it is more convenient to use matrix notation because of the large number of variables present. The true fluorescence decay can then be written in the following form:

$$\boldsymbol{F_N} = \boldsymbol{M}\boldsymbol{a}, \tag{15}$$

where $\boldsymbol{F_N}$ is the fluorescence decay binned in to $m$ channels, $\boldsymbol{a}$ is the vector (length $n$) of a-factor amplitudes, $\boldsymbol{a} = [a_1, a_2, a_3, \ldots, a_n]$, and $\boldsymbol{M}$, the decay matrix (of size $m \times n$), is defined as:

$$\boldsymbol{M_{ij}} = e^{-\frac{t_i}{\tau_j}}, \tag{16a}$$

$$\boldsymbol{M_{ij}} = e^{-k_j t_i}, \qquad (16b)$$

where $t_i$ is the $i^{\text{th}}$ channel (time point) and $\tau_j$ ($k_j$) is the $j^{\text{th}}$ lifetime (rate constant) of a pre-defined set (length $n$), which have corresponding a-factor amplitudes ($a_j$). Typically, the chosen set of lifetimes (rate constants) is logarithmically spaced to maximize the range of lifetimes (rate constants) while minimising the number of necessary components. Logarithmic spacing also facilitates the conversion (and thus comparison) between distributions of lifetimes and rate constants.

One of the challenges of using distributive methods for ill-posed problems is the vast number of degrees of freedom; there are many possible solutions and it is difficult for a standard least-squares analysis to optimize to a reliable minimum because of inherent instabilities. It is therefore necessary to regularize the problem by applying some additional constraint function, $C(\boldsymbol{a})$. The weighting of the constraint function is controlled by a regularisation parameter, $\gamma$, and the general mathematical description of the problem becomes of the form [3];

$$\dot{\boldsymbol{a}} = \arg\min_{a}\left[\frac{1}{2}\|\boldsymbol{F_N} - \boldsymbol{M}\boldsymbol{a}\|_2^2 + \gamma C(\boldsymbol{a})\right], \qquad (17)$$

where $\dot{\boldsymbol{a}}$ is the argument of the minimum (the set of values of $\boldsymbol{a}$ that minimises the function) and $\|\boldsymbol{F_N} - \boldsymbol{M}\boldsymbol{a}\|_2^2$ is the square of the Euclidean ($l_2$) norm of the difference between the decay, $\boldsymbol{F_N}$, and the fit, $\boldsymbol{Ma}$. Note that terms may contain some weighting factor that is not explicitly shown in the expression above.

*1.3.1. Maximum entropy method (MEM).* One approach to probabilistic analysis is the maximum entropy method (MEM), which has been used in a number of studies to analyse time-resolved fluorescence data [15-19]. In this case the constraint function is calculated as the Shannon-Jaynes entropy [15, 17];

$$C_{MEM}(\boldsymbol{a}) = -\sum_{j=1}^{n} a_j \log\left(\frac{a_j}{b_j}\right), \qquad (18)$$

where the set of values $\boldsymbol{b} = [b_1, b_2, b_3, \ldots, b_n]$ represent a default model for the system; however, as there is no default model that describes the dynamics of fluorescence emission, $b_j$ values are generally set to a constant value. Using constant $b_j$ favours equal contribution from all lifetimes and means that the introduction of structure into the distribution is discouraged [3]; in other words, only necessary components should be present in the solution. Note that (for the formalism used here) the regularisation parameter is defined to be negative ($\gamma < 0$) for the MEM; this reflects the fact that entropy should be maximised during the optimisation. Additionally, given that $\log(x)$ is only defined for $x > 0$, the MEM (as implemented here) inherently restricts probability amplitudes to positive values. It is worth pointing out that, while it is often physically justified to use strictly positive amplitudes, there are situations where negative amplitudes might be expected (such as in instances where solvent relaxation is required before fluorescence can occur, or in the case of excited state complex formation). Although not discussed further in this work, it should be noted that strategies have been developed to extend the MEM to include negative amplitudes [20-22].

*1.3.2. Smoothed exponential series method (SESM).* Siemiarczuk *et al.* [17] found little difference between the MEM and the exponential series method (ESM) [23, 24], which completely drops the constraint function (*i.e.* $C_{ESM} = 0$), except for in the case of a single-exponential decay (where iterative reconvolution would be the preferred method anyway). An initial attempt to replicate the ESM produced a discontinuous distribution of probability amplitudes. This did not seem physically realistic and so it was concluded that a *smoothness* constraint, $C_{SESM}$, should be applied. Phillips [25] showed that using the square of the second derivative was an effective strategy to constrain problems of this kind (an example of Tikhonov-Phillips regularisation, $C_{TP}(\boldsymbol{a}) = \|\boldsymbol{\Lambda a}\|_2^2$, with the Tikhonov matrix, $\boldsymbol{\Lambda}$, chosen to be the second derivative operator). The strategy used in this work follows the simplified implementation of Phillips smoothing used by Liu and Ware [26]; namely, the smoothness of the distribution was estimated by summing the square of the second difference between a-factor amplitudes;

$$C_{SESM}(\boldsymbol{a}) = \sum_{j=1}^{n-2}\left[\left(a_{j+2} - a_{j+1}\right) - \left(a_{j+1} - a_j\right)\right]^2. \qquad (19)$$

It is worth highlighting that, depending on the system being investigated, it may be beneficial to use other measures of smoothness for regularization. For example, a different order of derivative or, indeed, a combination of derivatives may provide a more appropriate constraint on the recovered amplitudes. Ideally, the chosen function should be directed by knowledge of the system since it will alter the efficacy of the regularization [27]. Here, the choice was primarily motivated by ease of computation.

*1.3.3. Basis pursuit denoising (BPDN).* Groma *et al.* [3] recently introduced an elegant approach to analysing time-resolved fluorescence data by using the $l_1$-norm of the vector of the *a*-factors as the constraint function;

$$C_{BPDN}(\boldsymbol{a}) = \|\boldsymbol{a}\|_1 = \sum_{j=1}^{n} |a_j|. \qquad (20)$$

In this case the formulation of the minimisation problem – equation (17) – becomes that of basis pursuit denoising (BPDN) [28, 29]. The analysis process for this

method can be summarised in the following way: find the simplest solution (by minimising $C_{BPDN}$), which can account for the experimental results (by minimising the least-squares error, $\chi_R^2$). This strategy follows the ideas of compressed sensing [30-32]; that it is possible to reconstruct the majority of a sparse (or compressible) signal by only using the most important elements.

*1.3.4. Other analysis methods.* This work only considers a sample of all possible analysis approaches and so it is perhaps useful to mention some other strategies that are available. For instance, lasso (least absolute shrinkage and selection operator) [33], which is closely related to BPDN, and its extension, elastic net [34], have previously been used to recover fluorescence lifetime distributions [35, 36]. The elastic net approach aims to improve lasso by combining it with Tikhonov-Phillips regularisation to achieve better performance with high-dimensional data with a small sample size, allow simultaneous selection of strongly correlated variables, and increase prediction accuracy [34].

The maximum likelihood method has also been successfully applied to fluorescence lifetime data [37, 38], and potentially performs better than least-squares analysis (as described in §1.2.1) with low count data by more accurately describing Poisson noise [37].

### *1.4. Content of paper*

The use of probabilistic methods in the context of fluorescence decay analysis is not a new concept; however, the value of this approach is not reflected within the literature, which is heavily dominated by iterative reconvolution analysis. There are a number of possible reasons for this. Firstly, iterative reconvolution is fast, versatile, reliable, and robust; such desirable attributes are extremely appealing for an analysis method. Secondly, commercial software packages almost exclusively use iterative reconvolution (mainly due to the former point); this has led to a familiarity, within the community, with this method and perhaps an inhibition of the use of less well-known methods. Thirdly, the results obtained from iterative reconvolution are intuitive; there is a direct link between fit parameters and model. The use of a well-defined model also means that it is (theoretically) easy to determine whether or not it is physically realistic. Finally, the computational power (speed) required for probabilistic methods has perhaps been lacking for "average" users until recently.

There is no doubt that iterative reconvolution is an effective analysis strategy for fluorescence decay data; however, as outlined above (and demonstrated below), it can suffer from deficiencies. Here we show that probabilistic analysis methods can complement iterative reconvolution analysis and help to overcome some of its limitations. In particular, there is focus on the benefit of gaining an unbiased assessment of the underlying decay dynamics. Since many may be unfamiliar with the probabilistic methods, a number of different approaches are evaluated with both simulated and experimental data. The results provide insight into the relative merits of the different methods and, at the same time, shed some light on the fluorescence behaviour of a complicated biophysical system; namely, 2-aminopurine (2AP) incorporated into nucleic acids.

The fluorescent base-analogue, 2AP, has become a valuable asset in the study of nucleic acids due to its highly desirable structural and photophysical properties; however, there still remains a significant shortfall in the understanding of the precise causes of its complex fluorescence behaviour within these systems. Of particular concern is the inability to provide a definitive physical model that can fully explain the multi-exponential fluorescence decay observed from time resolved fluorescence measurements of 2AP incorporated in oligonucleotides. Typically four-exponential terms are required to adequately fit the decays measured within the picosecond-nanosecond time range, which contrasts greatly with the single-exponential decay of free 2AP in solution. These four components are generally attributed to distinct conformational states in which 2AP experiences varying degrees of intermolecular interactions. Charge transfer between stacked bases has been implicated as the cause of the shortest lifetime ($\tau_1$, < 100 ps) observed. The longest lifetime ($\tau_4$, ~9 ns) is similar to that of free 2AP and is commonly accredited to a conformation where the 2AP moiety does not experience significant inter-base interaction. Such a state is thought to be accessible in DNA through base-flipping. The assignment of the two intermediate lifetimes (~0.5 ns and ~2 ns) has been more controversial, however, and there is still a lack of knowledge regarding their precise origin. It is possible that they are due to intermediate conformations between fully stacked and open forms, but alternative explanations, such as the existence of dark (non-fluorescent) states, have also been proposed. Indeed, rather than discrete components, a physically plausible explanation of these intermediate lifetimes is that they are, in fact, due to a broad distribution of decays that correspond to a whole range of conformations between stacked and unstacked extremes [6].

Based on redox potentials guanine (G) is predicted to efficiently quench excited-state 2AP; this behaviour is manifest in the significant amplitude of the short lifetime component of the fluorescence decay of 2AP when it is proximal to guanine within nucleic acids and in the dinucleotide 2AP-G. In contrast, inosine (I) is expected to be virtually redox-inactive with excited-state 2AP; however, a short lifetime component of 30 ps has been reported for the 2AP-I dinucleotide [39].

Dinucleotides offer a simple model of nucleic acids and have the significant benefit of limiting interactions to a single neighbouring base. In this work, the fluorescence decay profiles of 2AP-containing dinucleotides, 2AP-G

and 2AP-I, have been investigated to address the following questions. Is there evidence that the intermediate lifetimes are due to a broad distribution of states? Is there definitive evidence of the presence of the short lifetime component (implying charge transfer) when only inosine neighbours 2AP?

## 2. Experimental / Methods

### *2.1. Experimental and simulated decays*

Simulated decays were created by convolution of a known decay function with an experimental IRF (Supplementary Information, figure S1). The same IRF was used during fitting of these decays. Poisson noise was added after convolution to avoid imposing IRF structure into the noise. The number of counts in the peak channel was typically set to around 10,000 to match the experimental data (see below). A summary of the simulated decays is given in table 1.

One of the simulated decays was based on a Γ distribution of lifetimes to test the ability of the analysis techniques to handle non-exponential decays. In this case the probability for lifetime $\tau$ (decay constant $k$) is given by;

$$p(\tau) = \frac{\beta^{\phi}}{\Gamma(\phi)}\left(\frac{1}{\tau}\right)^{\phi+1} e^{-\frac{\beta}{\tau}}, \qquad (21a)$$

$$p(k) = \frac{\beta^{\phi}}{\Gamma(\phi)} k^{\phi-1} e^{-\beta k}, \qquad (21b)$$

where $\phi$ and $\beta$ are shape and scale parameters, respectively, and $\Gamma(x)$ is the Gamma function. An analytical summary of the properties of the Γ distribution is given by Fogarty *et al.* [6]. Of relevance to the present study is the fact that the decay function when using the Γ distribution is relatively simple and that the distribution of lifetimes (rate constants) has a well-defined mode,

$$\tau_j^* = \frac{\beta_j}{(\phi_j + 1)}, \qquad (22a)$$

$$k_j^* = \frac{\phi_j - 1}{\beta_j}, \qquad (22b)$$

and mean,

$$\langle\tau\rangle_j = \frac{\beta_j}{\phi_j - 1}, \qquad (23a)$$

$$\langle k\rangle_j = \frac{\phi_j}{\beta_j}, \qquad (23b)$$

which correspond to the most probable lifetime (22a) or rate constant (22b) and the average lifetime (23a) or rate constant (23b), respectively.

It is worth pointing out that the Gamma distribution described above is equivalent to the 'Becquerel function' introduced by Berberan-Santos and coworkers [40] with appropriate substitution ($\beta = \tau_0/c$ and $\phi = 1/c$). An important outcome from this previous work is the ability of the distribution-based decay function to describe complex photophysical systems, possibly in a more physically realistic way than would be feasible with a conventional multi-exponential analysis.

*2.1.1. Time-correlated single photon counting (TCSPC) experiments.* TCSPC was used to measure the fluorescence decay profiles of 2AP-G and 2AP-I. Samples were measured in fused silica cells (Starna) which had a 1 cm path length. Fluorescence decay curves were recorded using an Edinburgh Instruments spectrometer (FL920) equipped with TCC900 photon counting electronics. The excitation source was a tuneable, mode-locked Ti:Sapphire laser system (Coherent MIRA Ti:Sapphire laser pumped by a Coherent 10 W Verdi CW laser), producing approximately 200 fs pulses at a repetition rate of

**Table 1.** Decay functions and parameters used for simulated data.

| Decay Type | Code | Decay Function | Parameters | | | |
|---|---|---|---|---|---|---|
| Single-exponential | 1δ | $F_N = a_j e^{-\frac{t}{\tau_j}}$ | **$j$** | **$a_j$** | **$\tau_j$/ns** | |
| | | | 1 | 1.00 | 5.00 | |
| Four-exponential | 4δ | $F_N = \sum_{j=1}^{4} a_j e^{-\frac{t}{\tau_j}}$ | **$j$** | **$a_j$** | **$\tau_j$/ns** | |
| | | | 1 | 0.08 | 0.06 | |
| | | | 2 | 0.32 | 0.67 | |
| | | | 3 | 0.41 | 3.03 | |
| | | | 4 | 0.19 | 11.11 | |
| Gamma distribution (Gamma1) | 1Γ | $F_N = \frac{a_j}{(1 + t/\beta_j)^{\phi_j}}$ | **$j$** | **$a_j$** | **$\beta_j$/ns** | **$\phi_j$** |
| | | | 1 | 1.00 | 100 | 21.0 |

76 MHz. A pulse picker (Coherent 9200) was used to reduce the pulse repetition rate to 4.75 MHz. A harmonic generator (Coherent 5-050) was used to triple the frequency of the source light. Fluorescence emission was detected orthogonal to the excitation beam through a polariser set at the magic angle (54.7°) with respect to the vertically polarised excitation. A band-pass of 18 nm was used in the emission monochromator and photons were detected using a cooled microchannel plate detector (Hamamatsu R3809 series). All experiments were performed at an excitation wavelength of 305 nm. A 320 nm long pass filter Schott Glass (Newport) was used to block scattered light during fluorescence decay measurements. Fluorescence decay curves were recorded on a time scale of 50 ns, binned into 4096 channels, to a total of 10,000 counts in the peak channel. Decays were measured at an emission wavelength of 380 nm. The instrument response function was recorded using scattered light from Ludox solution at 305 nm. All measurements were made at room temperature (20°C).

2AP-containing dinucleotides were purchased from ATDBio Ltd. Samples were dissolved in Tris buffer (containing 0.02 M Tris-HCl pH 7.5 and 0.1 M NaCl) at a concentration of approximately 10 μM.

*2.1.2. Decay fitting.* All fits were performed on a standard PC using in-house scripts written for MATLAB (R2015a, The MathWorks, Inc., Natick, Massachusetts, United States of America.). The salient features of each analysis technique will be addressed below.

During fitting the background level was fixed to a predetermined value based on the data before the IRF rise. The fitting range was generally defined to start from the peak channel of the IRF and end at the channel of the fluorescence decay that was either around 10 times higher than the background level or equivalent to 0.1% of the counts in the peak channel, whichever was greater.

Fitting strategies for the iterative reconvolution analysis used in this work were partly inspired by Decay Fit (FluorTools, www.fluortools.com). The use of in-house scripts allowed the inclusion of non-exponential decay models, such as the Gamma distribution described above.

For the smoothed exponential series method (SESM) and the maximum entropy method (MEM) a total of 201 fixed lifetimes, logarithmically spaced between 0.01 ns and 100 ns, were used during optimisation. The a-factors were restricted to values between 0 and 1. Since its associated constraint function is only defined for positive, non-zero a-factors, the lower limit for these in the MEM analyses was set to $10^{-10}$. In attempting to avoid local minima, five different starting points were used; four homogeneous distributions where all a-factors were initially set to the same constant ($1/n$ (where $n$ = number of a-factors), 0, 0.5, or 1) and one distribution where a-factors were initially set to randomly generated values. To prevent fitting noise, a-factors were initially set (but not fixed) to their lower limit if the corresponding lifetime was on the same order-of-magnitude as the channel width (0.0122 ns). The optimisation which gave the best minimisation out of the five trials was further optimised to provide the fit model. The regularisation parameter, $\gamma$, was set to a value of 0.1 for the SESM and $-10^{-3}$ for the MEM. These values were found (by trial-and-error) to be a good trade-off between improving the smoothness (continuity) of the distribution while maintaining a reasonable resolution in the lifetime dimension, which allowed discrete lifetimes to be recovered.

The SparseLab toolbox [41] from Stanford was used for optimisation during basis pursuit denoising (BPDN, compressed sensing) analysis. Groma *et al.* investigated a number of different minimisation algorithms for BPDN but found that the one used by SparseLab (primal-dual log-barrier algorithm [29]) was best suited for fitting fluorescence decays [3]. The regularisation parameter, $\gamma$, was set to a value of $10^{-3}$ for BPDN analysis. This value was found to give a good balance between finding a sparse solution while avoiding oversimplification of the underlying dynamics. A total of 601 fixed lifetimes, logarithmically spaced between 0.01 ns and 100 ns, were used during the optimisation.

*2.1.3. Instrument response function shift.* For iterative reconvolution (ItRe) analysis, the temporal shift between the IRF and measured decay was optimised as an additional parameter during the fitting process. In contrast, the shift was fixed during optimisation of the amplitudes of the SESM, the MEM, and BPDN analysis methods. This was done partly to circumvent the need to continually convolute the large decay matrix with a shifting IRF. It also avoided the potential issue of an erroneously large shift being compensated by short lifetimes on the order of the channel width. An iterative approach was therefore taken to ensure the best shift was obtained during fitting with the probabilistic methods. After optimising the amplitudes with a fixed shift, the shift was optimised with fixed amplitudes. This process was repeated until there was an insignificant change in the quality of fit. The final temporal shift was generally found to agree within a fraction of a channel for all methods using this approach. For the simulated data, all methods typically gave shift values within $1/20^{th}$ of a channel from the expected shift value of zero. For the experimental data, the maximum discrepancy between methods was around $1/3^{rd}$ of a channel.

The iterative approach to the shift optimisation was somewhat unavoidable for the BPDN analysis due to the use of the SparseLab toolbox (though the open-source nature of this toolbox could allow future development to resolve this issue). Although the approach was perhaps not ideal, the optimisation of the temporal shift was found to be especially important for BPDN analysis. Early

simulations without such an optimisation (not shown) often showed the presence of erroneous negative amplitudes at short lifetimes (presumably compensating for the poorly characterised shift parameter).

In the interest of open access, the code that has been written to perform the analysis described in this paper is available to download from a data share repository (see details in the Acknowledgements). Note that the BPDN analysis requires additional files from the SparseLab toolbox which will not be included here but can be freely accessed elsewhere [41].

## 3. Results

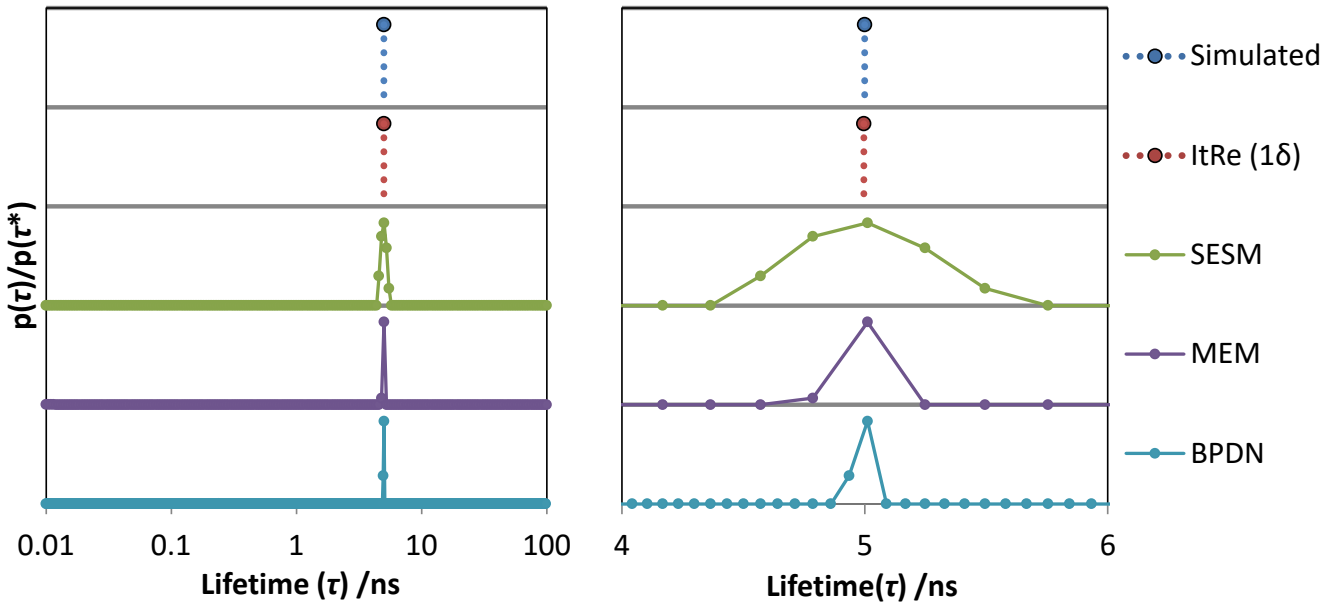
Comparison of the fitting methods is most easily achieved by visual inspection. For each system, probability amplitudes, normalized by their mode, are plotted against their corresponding lifetime on a logarithmic abscissa; an expanded view on a linear abscissa is also shown. Discrete exponentials are represented as circular markers (joined by dotted lines that are physically meaningless, but are presented to guide the eye), while distributions are shown as continuous curves with solid markers.

The peaks in the distributions of the SESM and the MEM analyses were characterised by fitting Gaussian functions to obtain a-factors (defined by peak area) and full-width at half-maximum (FWHM) for each decay component; this required transformation of the probability amplitudes to account for the use of logarithmically spaced lifetimes [17]. In practice, recovered amplitudes were divided by their associated lifetime to compensate for fitting linearly spaced Gaussian distribution functions to logarithmically spaced peak amplitudes. (Note that all plots show untransformed $a$-factor amplitudes that have been normalized and not the Gaussian fits.) The FWHM value reflects the spread of lifetimes that accounts for the decay component. Note that FWHM values are not given in the case of the BPDN analysis because, typically, only a single value contributed to the decay component. On occasion the peaks in the BPDN analysis were composed of two (or three) neighbouring components. In this case the amplitude-weighted lifetime, $\langle\tau\rangle = \sum_j^n a_j\tau_j$, of the contributing components and the sum of their a-factors is reported as the associated lifetime and "area" (a-factor), respectively. The amplitude-weighted lifetime of the overall decay is proportional to the integrated area under the associated decay curve and so is provided as a simple measure of the character of the calculated decay model.

Due to the diversity of the different analytical methods used during the study, it was necessary to devise a standard measure of the quality of fit. As the degrees of freedom in the probabilistic analysis methods was sometimes ambiguous (for example, in the case of the BPDN analysis a total of 601 amplitudes could vary; however, the final number of parameters that contributed significantly to the fit was typically less than four and, in addition, the parameters were not fully independent of each other), the $\chi^2$ value is reported in addition to $\chi^2_R$. The number of fitted channels, $N_c$, is provided in the table captions as a marker of the "ideal" $\chi^2$ value for the fit.

It should be assumed that, unless specifically stated, all the fits reported were of adequate quality to be considered a fair representation of the data. For reference, all fits and associated data (*e.g.* residuals and fitting parameters) can

**Figure 1.** Comparison of the results of fitting a simulated single-exponential decay (the parameters of which are shown in the top panel). Left: Logarithmic abscissa. Right: Linear abscissa between 4 ns and 6 ns.



be found in the summary files supplied in the Supplementary Information.

*3.1. Simulated decays*

*3.1.1. Single-exponential function.* Figure 1 shows a visual representation of the fits obtained for the simulated single-exponential decay (table 1) and table 2 shows the associated fitted parameters. This example represents the simplest possible fluorescence decay system; the decay is defined by a single lifetime, 5 ns, that is considerably longer than the IRF width (~100 ps) but only a fraction of the time range of the measurement (~50 ns). In addition, only Poisson noise is present in the simulated decay and there is also no need to compensate for temporal shift in the IRF position. If an analysis approach is to have any standing then it must be able to accurately describe this system.

Encouragingly, the results show that all methods were successful when analysing the simulated single-exponential decay. All methods show a single component at a lifetime close to the simulated value of 5 ns. Distributive methods have negligible probability amplitudes at other lifetime values. The fitted lifetime has less than 1% deviation from the simulated value and $\chi^2$ values are generally consistent between all methods suggesting similar qualities of fit.

The SESM and MEM distribution fits both give fairly narrow peaks. Ideally the distributive methods would give a single point but this is unreasonable to expect, given the presence of noise and limitations of the optimisation. Consistent with previous observations [17], the MEM performs slightly better than the SESM in terms of peak width. Note that the peaks could have been narrowed by decreasing the weight of the regularisation parameter; however, it was deemed more appropriate to use consistent regularisation parameter for each analysis method across all of the systems studied. Using a consistent value ensured that any difference observed in peak widths between systems were due to differences in the underlying decay distribution, rather than being due to the choice of regularisation parameter. The influence of noise is most significant when considering times on the same order of magnitude as the channel width (~0.01 ns) and so it is reassuring to see that, despite being adjustable parameters, the distributive methods show negligible amplitude at short lifetime values.

**Table 2.** Summary of lifetimes obtained for fits to a simulated single-exponential decay ($N_c$: 2412).

| 1δ | $\tau$ (FWHM)/ns | $\chi_R^2$ ($\chi^2$) |
|---|---|---|
| Simulated | 5.00 | - |
| ItRe (1δ) | 5.00 | 0.98 (2365) |
| SESM | 4.99 (0.65) | 0.99 (2394) |
| MEM | 5.01 (0.17) | 0.98 (2366) |
| BPDN | 5.01 | 0.99 (2376) |

**Figure 2.** Comparison of the results of fitting a simulated four-exponential decay (the parameters of which are shown in the top panel). Left: Logarithmic abscissa. Right: Linear abscissa up to 15 ns.

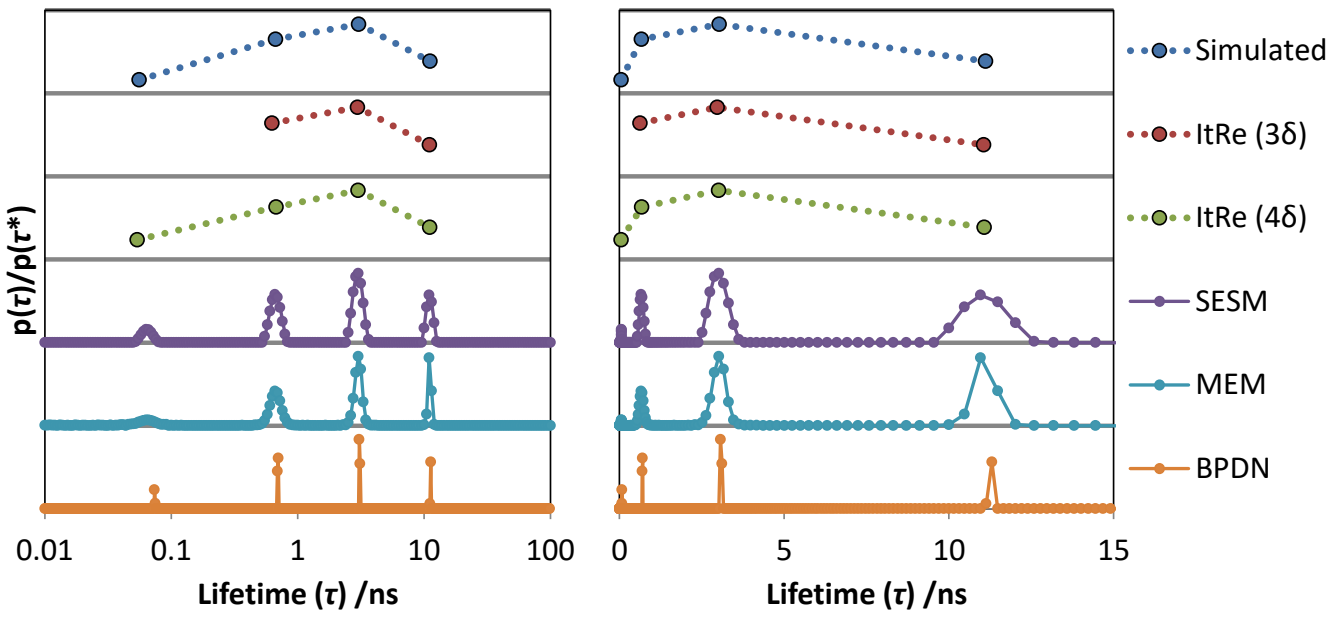
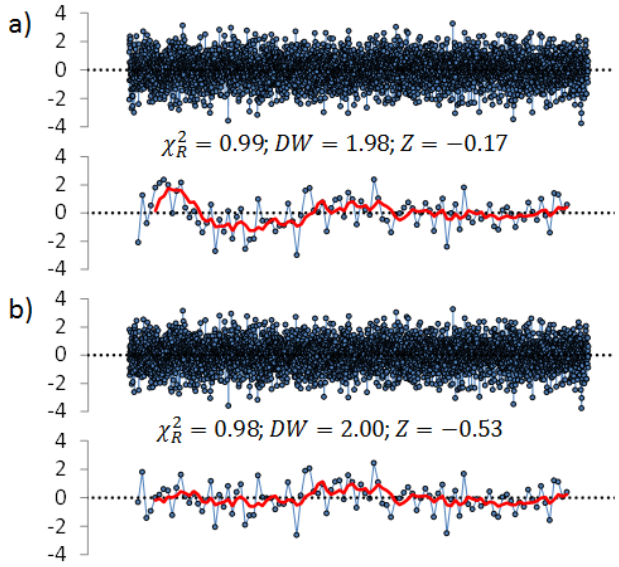


**Table 3.** Summary of lifetime parameters obtained for fits to a simulated four-exponential decay ($N_c$: 3567).

| 4δ | $\tau$ (FWHM)/ns | | | | a-factor | | | | $\langle\tau\rangle$/ns | $\chi_R^2$ ($\chi^2$) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau_1$ | $\tau_2$ | $\tau_3$ | $\tau_4$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | | |
| Simulated | 0.06 | 0.67 | 3.03 | 11.11 | 0.08 | 0.32 | 0.41 | 0.19 | 3.57 | - |
| ItRe (3δ) | - | 0.63 | 2.97 | 11.05 | - | 0.35 | 0.45 | 0.21 | 3.82 | 0.99 (3525) |
| ItRe (4δ) | 0.05 | 0.67 | 3.01 | 11.07 | 0.11 | 0.30 | 0.40 | 0.19 | 3.46 | 0.98 (3492) |
| SESM | 0.06 (0.01) | 0.67 (0.15) | 3.00 (0.62) | 11.06 (1.57) | 0.08 | 0.31 | 0.41 | 0.19 | 3.59 | 0.98 (3496) |
| MEM | 0.07 (0.02) | 0.67 (0.15) | 3.03 (0.47) | 11.09 (0.76) | 0.08 | 0.30 | 0.41 | 0.19 | 3.52 | 0.98 (3496) |
| BPDN | 0.07 | 0.70 | 3.09 | 11.29 | 0.09 | 0.32 | 0.41 | 0.19 | 3.59 | 0.98 (3509) |

**Figure 3.** Residuals associated with (a) the 3-exponential fit and (b) the 4-exponential fit of the simulated 4-exponential decay. Upper plots show residuals for all fitted channels (~40 ns) while the lower plot shows residuals for only the first 100 points (~1 ns) of the fit. The red line in the lower plots is a moving average fit of the residuals based on 5 consecutive points. The $\chi_R^2$, Durbin-Watson ($DW$), and Runs Test statistic ($Z$) for each fit is also given.




*3.1.2 Four-exponential function.* Figure 2 shows a visual representation of the fits obtained for the simulated four-exponential decay (table 1) and table 3 shows the associated fitted parameters. This example was used to determine how well the analysis methods would cope with the presence of multiple decay components that were distributed over a large timescale range. The lifetime values used were inspired by the fluorescence decay observed for 2AP-containing dinucleotides (see below) and so this simulation can be taken as an idealised version of an experimental system.

Generally speaking, the analysis methods perform very well at recovering the underlying model. All probabilistic methods (SESM, MEM, and BPDN) show four distinct components. Lifetimes and probability amplitudes are visually well-matched to the simulated decay parameters. Given the complexity of the simulated decay, it is particularly pleasing to see the success of the distributive methods, which were not restricted by any functional form during fitting.

This example provides useful insight into the challenge of analysing time-resolved data. Figure 3 shows residuals associated with the three- (3δ) and four-exponential (4δ) fits of the decay. Superficially, the residuals of the 3δ fit appear to be randomly distributed (figure 3a, upper plot) and the associated statistical parameters ($\chi_R^2$, $DW$, and $Z$) are acceptable (indeed, close to ideal values; 1, 2, and 0, respectively); this all suggests that the model is a fair representation of the data (*i.e.* it is a good fit). In fact, on the basis of the multi-exponential fits, there would appear to be no reason to believe a fourth component is necessary. This perfectly illustrates the importance of ensuring that the residuals are properly inspected at **all** points in the fit. The lower plots in figure 3 show the first 100 residuals (covering around 1 ns of the initial decay). On this timescale it is clear that the three-exponential model does not describe the data as well as the four-exponential model; there is an obvious oscillation in the residuals (highlighted by the moving average fit shown in red). Without the knowledge that the simulated decay was produced with four exponentials, it would have been quite easy to miss this deviation at early times. This brings to light an important benefit of the probabilistic methods;

**Figure 4.** Comparison of the results of fitting a simulated 1Γ decay (the parameters of which are shown in the top panel). Left: Logarithmic abscissa. Right: Linear abscissa up to 12 ns.

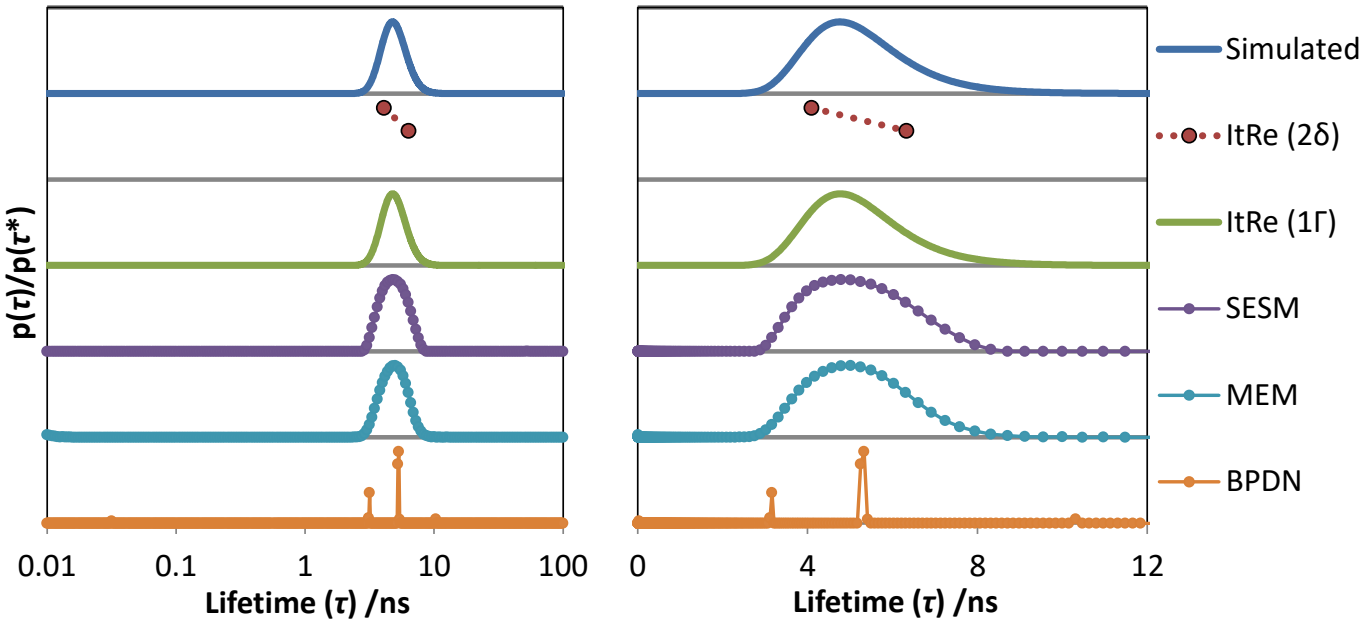



**Table 4.** Summary of fit parameters obtained from analyses of a simulated 1Γ decay ($N_c$: 2610).

| Gamma1 | $\tau$ (FWHM)/ns | | | a-factor | | | $\langle\tau\rangle$/ns | $\chi^2_R$ ($\chi^2$) |
|---|---|---|---|---|---|---|---|---|
| | $\tau_1$ | $\tau_2$ | $\tau_3$ | $a_1$ | $a_2$ | $a_3$ | | |
| Simulated | $\langle\tau_\Gamma\rangle$: 5.00 $\beta_\Gamma$: 100.00 $\tau_\Gamma^*$: 4.55 $\phi_\Gamma$: 21.00 | | | | $a_\Gamma$: 1.00 | | 5.00 | - |
| ItRe (2δ) | 4.09 | - | 6.33 | 0.60 | - | 0.40 | 5.00 | 0.97 (2536) |
| ItRe (1Γ) | $\langle\tau_\Gamma\rangle$: 5.00 $\beta_\Gamma$: 104.31 $\tau_\Gamma^*$: 4.56 $\phi_\Gamma$: 21.86 | | | | $a_\Gamma$: 1.00 | | 5.00 | 0.97 (2538) |
| SESM | - | 5.17 (2.56) | - | - | 1.00 | - | 5.17 | 0.97 (2535) |
| MEM | - | 5.09 (2.51) | - | - | 1.00 | - | 5.09 | 0.97 (2535) |
| BPDN | 3.16 | 5.30 | 10.31 | 0.21 | 0.77 | 0.02 | 4.98 | 0.98 (2551) |

namely, they all correctly predicted the presence of the fourth, short lifetime component.

*3.1.3. Gamma distribution.* Figure 4 shows a visual representation of the fits obtained for a simulated decay based on a Γ distribution of lifetimes (Gamma1, table 1) and table 4 shows the associated fitted parameters. This example was created to establish whether or not the analysis methods were capable of modelling a system where there was a spread of similar lifetimes rather than well-defined, discrete lifetimes. This situation might arise, for example, in a Förster resonance energy transfer (FRET) system where there is a distribution of distances between donor-acceptor chromophore pairs [6, 17, 42, 43].

The fitting results for this system exemplify the difficulty in recovering the true underlying model when there is no *a priori* knowledge to inform judgement. The different fitting methods delivered different results, but all of the fitted models adequately describe the measured decay (under the chosen criteria) and so it would be impossible to determine the true model by a simple comparison of the quality of the fits. Given the disparity of the models, it is particularly unsettling that the 2δ and 1Γ fits appear to describe the data equally well.

The precise value of the fit parameters is not of vital importance in the message provided by this example. The salient feature of the analysis is that the methods complement each other and can be used to guide the experimentalist to the most rational choice of model. The probabilistic approaches provide a general consensus that favours a model based on a distribution of lifetimes rather than discrete lifetimes. For instance, the FWHM of the peaks in the SESM and the MEM fits is significantly greater than that obtained for the discrete components analysed in the previous examples. This gives weight to the idea of a distribution (or at least unresolvable components) being responsible for the observed decay in this case. Although the BPDN analysis approach is not well-suited to uncovering a distribution of decays [3], the shape of the BPDN amplitudes clearly correlates with the results from the SESM and the MEM analyses. The distribution of amplitudes is also consistent with that expected for a multi-exponential fit of a broad distribution of lifetimes; namely, there is a major peak at the distribution maximum and two other peaks approximately $\pm\sigma$ from this position [17] (note the Gamma function is asymmetric, which accounts for the relatively small amplitude component on the long lifetime side of the major peak). Based on the insight obtained from the probabilistic methods, the iterative reconvolution method, using the gamma function, (ItRe (1Γ)) can then be recognised to provide a more accurate description of the lifetime distribution; as is seen in the well-matched parameters obtained.

*3.1.4 Two-exponential function based on ItRe (2δ) fit to gamma distribution.* One way to improve upon the confidence of a particular model is to simulate a decay based on the calculated parameters and then re-fit to see if a consistent set of probability amplitudes is obtained [17]. Figure 5 shows a visual representation of the fits obtained for a two-exponential decay simulated using the parameters from the discrete exponential fit of Gamma1. This decay simulation will be denoted as Exp2G. Table 5 shows the associated fitted parameters from this system.

As expected, there is very little difference between the parameters obtained for the 2δ fit and those input to the simulated decay. However, the Γ distribution fit also adequately describes the two-exponential decay data. Indeed, there is great similarity between the original Γ distribution (shown by the dashed, black line in figure 5) and the Γ fit for the two-exponential decay. This means that, at the signal-to-noise ratio used ($10^2$), iterative reconvolution alone cannot bias the choice of model underlying Gamma1 and Exp2G decays; in each case, 2δ and 1Γ models are equally acceptable candidates. Without gaining further insight, it would be tempting to assign the Γ distribution model to both Gamma1 and Exp2G simulations since it is simpler (*i.e.* has fewer adjustable parameters) than the 2δ model.

**Figure 5.** Comparison of the results of fitting a simulated two-exponential decay based on the parameters of the ItRe (2δ) fit of the 1Γ decay. Left: Logarithmic abscissa. Right: Linear abscissa up to 12 ns. The parameters of the simulated decay are shown in the top panel, together with the original Γ distribution, shown by the dashed, black line.

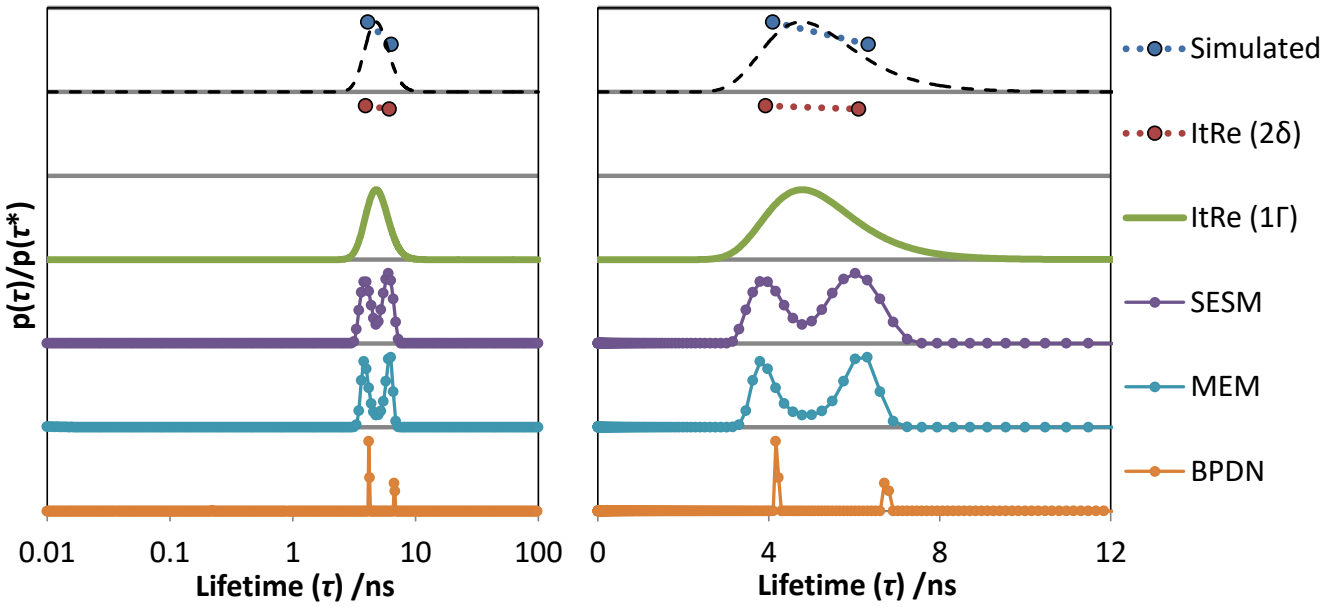


**Table 5.** Summary of lifetime parameters obtained for fits to a simulated two-exponential decay ($N_c$: 2561) based on the parameters of ItRe (2δ) fit to the 1Γ decay given in table 4.

| Exp2G | $\tau$ (FWHM)/ns | | a-factor | | $\langle\tau\rangle$/ns | $\chi_R^2$ ($\chi^2$) |
|---|---|---|---|---|---|---|
| | $\tau_1$ | $\tau_2$ | $a_1$ | $a_2$ | | |
| Simulated | 4.09 | 6.33 | 0.60 | 0.40 | 5.00 | - |
| ItRe (1Γ) | $\langle\tau_\Gamma\rangle$: 5.00 $\tau_\Gamma^*$: 4.58 | $\beta_\Gamma$: 108.98 $\phi_\Gamma$: 22.79 | $a_\Gamma$: 1.00 | | 5.00 | 0.98 (2586) |
| ItRe (2δ) | 3.93 | 6.10 | 0.51 | 0.49 | 4.99 | 0.97 (2561) |
| SESM | 3.94 (0.88) | 5.99 (1.50) | 0.46 | 0.54 | 5.05 | 0.97 (2563) |
| MEM | 3.91 (0.68) | 6.11 (1.09) | 0.48 | 0.52 | 5.05 | 0.97 (2563) |
| BPDN | 4.19 | 6.75 | 0.68 | 0.32 | 5.00 | 0.97 (2575) |

Importantly, there is a significant difference in the results obtained for the distributive methods; all of these methods show two well-defined components rather than a single broad peak (SESM and MEM) or three discrete components (BPDN) that was observed for the Gamma1 decay. This discrepancy helps to rule out the 2δ model as a plausible description of the Gamma1 system (which is, of course, consistent with the fact that it is the wrong model). If the 2δ model had been correct then the Gamma1 and Exp2G decays would have been equivalent, within error, from the perspective of the probabilistic methods, thus, the distributions obtained should have been similar.

Note that there is still no guarantee about any of the models being correct, the Exp2G simulation simply shows that a 2δ model is unlikely to be the true model for the Gamma1 system. Nevertheless, careful consideration of the results from all of the different methods would hopefully lead to the conclusion that the Gamma1 decay was due to a 1Γ model (or at least a model based on a distribution of decay lifetimes) and the Exp2G decay was due to a 2δ model.

This example highlights the potential power of using probabilistic methods in combination with deterministic ones. Without imposing a model form, the distributive methods found a broad peak for the simulated Γ distribution but found a double peak for the simulated 2δ system. This shows that probabilistic methods could be used to give an unbiased estimate of the underlying model. On the other hand, the deterministic (iterative reconvolution) method provided more accurate model parameters than the probabilistic approaches when the true underlying model was used. Deterministic methods could therefore be used to refine the model that was initially determined by a probabilistic approach.

Of course, it is easy to argue for a particular model when the true model is already known. It may be useful to perform blind simulations in the future to determine the practical reliability of combining analysis methods. This being said, while simulated decays provide great insight into the different analysis strategies, the capability of an analysis method should really be assessed with real, experimental data, which may contain unexpected distortions that are not considered in simulations.

### *3.2. Experimental decays*

*3.2.1. 2AP-G dinucleotide.* Figure 6 shows a visual representation of the fits obtained for the fluorescence decay of the 2AP-G dinucleotide and table 6 shows the associated fitted parameters.

The analysis methods generally show four components that are consistent with the typical parameters obtained for 2AP in nucleic acid constructs [39, 44]. The similarity of the fit parameters obtained for the different methods gives confidence in the results. The probabilistic methods also give some insight into the underlying dynamics responsible for the fluorescence decay. The results presented here favour the more conventional interpretation of the intermediate decay components as two distinct lifetimes. All of the methods exhibit well-separated components and show no indication of a broad distribution being responsible for the intermediate components. Furthermore, the widths of the peaks obtained from the probabilistic methods are comparable to those obtained for the simulated decays based on discrete components and far smaller than those observed for the Gamma1 lifetime distribution.

**Figure 6.** Comparison of fits for the fluorescence decay of a 2AP-G dinucleotide (excitation and emission wavelengths were 300 nm and 380 nm, respectively). Left: Logarithmic abscissa. Right: Linear abscissa up to 12 ns.

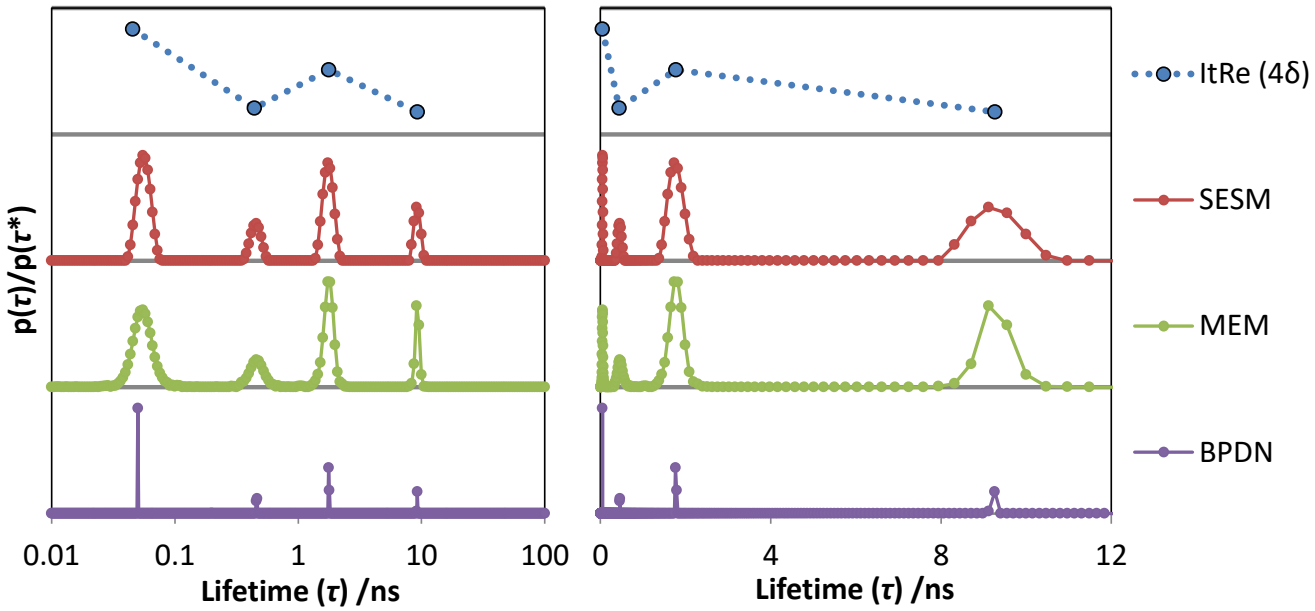



**Table 6.** Summary of lifetime parameters obtained for fits to the fluorescence decay of a 2AP-G dinucleotide ($N_c$: 3090).

| 2AP-G | $\tau$ (FWHM)/ns | | | | a-factor | | | | $\langle\tau\rangle$/ns | $\chi_R^2$ ($\chi^2$) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau_1$ | $\tau_2$ | $\tau_3$ | $\tau_4$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | | |
| ItRe (4δ) | 0.05 | 0.44 | 1.77 | 9.27 | 0.48 | 0.12 | 0.30 | 0.10 | 1.55 | 0.99 (3060) |
| SESM | 0.06 (0.16) | 0.46 (0.10) | 1.76 (0.41) | 9.24 (1.38) | 0.44 | 0.12 | 0.33 | 0.11 | 1.70 | 0.99 (3070) |
| MEM | 0.06 (0.02) | 0.47 (0.14) | 1.78 (0.37) | 9.27 (0.82) | 0.44 | 0.12 | 0.33 | 0.11 | 1.70 | 0.99 (3070) |
| BPDN | 0.05 | 0.46 | 1.77 | 9.24 | 0.47 | 0.12 | 0.30 | 0.11 | 1.61 | 1.00 (3083) |

*3.2.2. 2AP-I dinucleotide.* Figure 7 shows a visual representation of the fits obtained for the fluorescence decay of the 2AP-I dinucleotide and table 7 shows the associated fitted parameters.

Again, the methods generally show four components that are consistent with the typical parameters obtained for 2AP in nucleic acid constructs. The 2AP-G and 2AP-I results differ significantly in the contribution from the short lifetime component; indeed, it is questionable whether the component is real or spurious for 2AP-I. The amplitude of the short lifetime component is very small and this is compounded by the fact that the nominal lifetime value varies considerably between the different analysis methods.

Figure 8 shows residuals associated with the three- and four-exponential fits of the 2AP-I decay. Similar to the simulated four-exponential decay results, the 3δ fit residuals appear to be randomly distributed (figure 8a, upper plot) and the associated statistical parameters ($\chi_R^2$, *DW*, and *Z*) are acceptable suggesting that the model is a fair representation of the data. Unlike the simulated four-exponential decay results, the addition of another component makes very little difference to the residuals. There is slight evidence of improvement on the short timescale (figure 8, lower plots), but the difference is almost negligible compared to the random variation observed at longer times. It is difficult to argue the case for the need of a fourth exponential term. Nevertheless, the probabilistic methods do all exhibit a short lifetime peak, albeit a very small (~2%) component (note that this peak in the MEM distribution has very low amplitude, due to its broad width; the peak is essentially invisible in the plot in figure 7). Although it would be naïve to present

**Figure 7.** Comparison of fits obtained for the fluorescence decay of a 2AP-I dinucleotide (excitation and emission wavelengths were 300 nm and 380 nm, respectively). Left: Logarithmic abscissa. Right: Linear abscissa up to 12 ns.

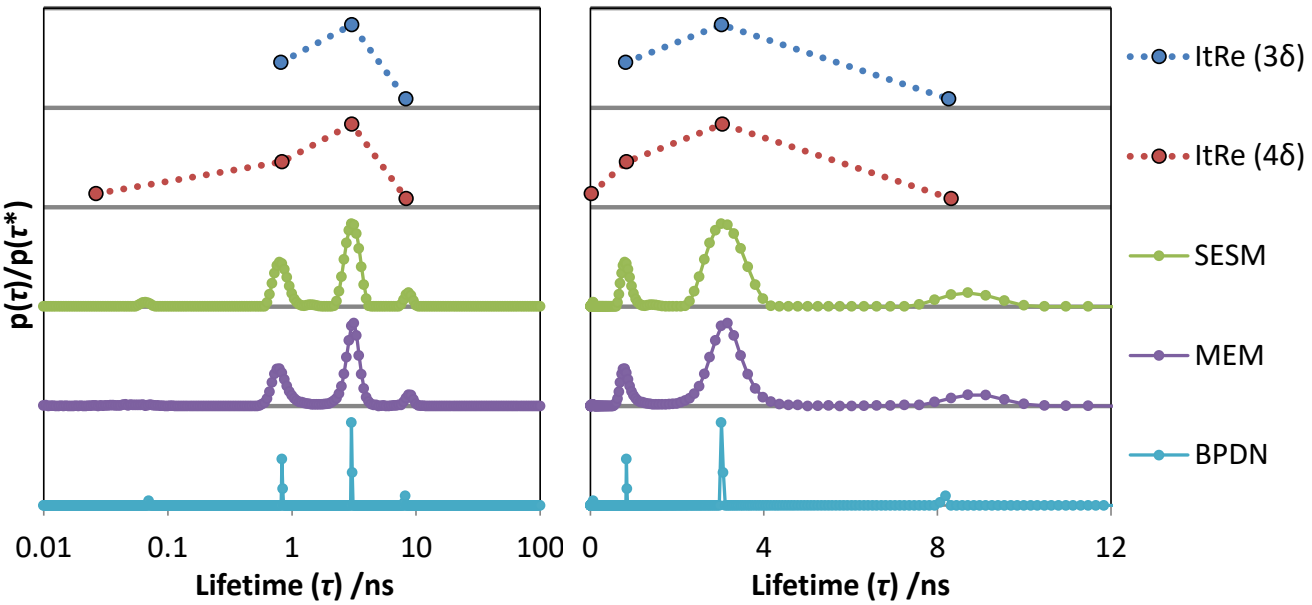



**Table 7.** Summary of lifetime parameters obtained for 2AP-I dinucleotide fluorescence decay ($N_c$: 2097).

| 2AP-I | $\tau$ (FWHM)/ns | | | | a-factor | | | | $\langle\tau\rangle$/ns | $\chi_R^2$ ($\chi^2$) |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\tau_1$ | $\tau_2$ | $\tau_3$ | $\tau_4$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | | |
| ItRe (3δ) | - | 0.82 | 3.03 | 8.27 | - | 0.33 | 0.61 | 0.06 | 2.63 | 1.06 (2215) |
| ItRe (4δ) | 0.03 | 0.83 | 3.04 | 8.32 | 0.09 | 0.30 | 0.55 | 0.06 | 2.41 | 1.06 (2211) |
| SESM | 0.07 (0.01) | 0.82 (0.23) | 3.10 (0.93) | 8.70 (1.47) | 0.02 | 0.31 | 0.61 | 0.05 | 2.63 | 1.05 (2211) |
| MEM | 0.05 (0.03) | 0.80 (0.24) | 3.13 (0.89) | 8.87 (1.55) | 0.02 | 0.30 | 0.62 | 0.05 | 2.65 | 1.05 (2212) |
| BPDN | 0.07 | 0.84 | 3.03 | 8.16 | 0.03 | 0.32 | 0.59 | 0.06 | 2.57 | 1.06 (2229) |

these results as definitive evidence of a fourth component in the fluorescence decay of 2AP-I, they show that such a component can be detected in an unbiased optimisation. For a more conclusive analysis it would be necessary to improve the quality of the data (for example, by increasing the counts in the peak channel) or perhaps perform global analysis on multiple decays. It may also be necessary to use a technique with higher temporal resolution (such as fluorescence upconversion) to ensure the observation of a short lifetime component in 2AP-I is not simply an artefact of shortcomings in convolution strategy (*e.g.* imperfect IRF shift optimisation). Again, the precise detail of the analysis presented here is not the main message of this contribution. The important consideration is that probabilistic methods can provide insight into the underlying dynamics that might otherwise be overlooked.

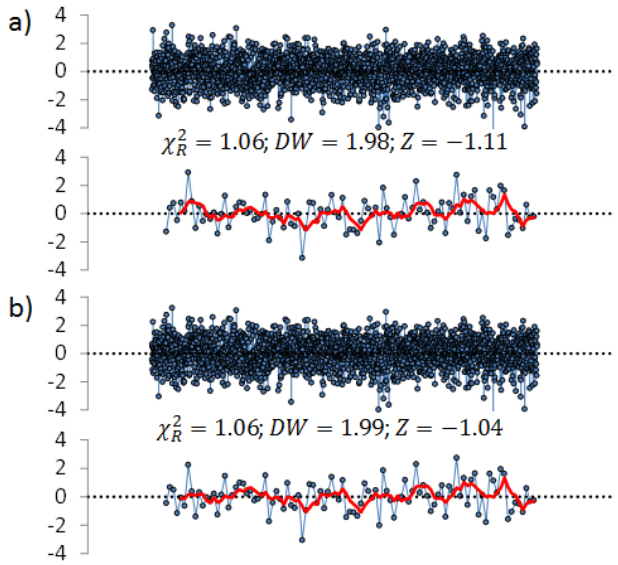
## 4. Discussion

In light of the results presented above, benefits and limitations of the various analysis methods that were used during this study will now be reviewed.

### *4.1. Iterative reconvolution*

As expected, the iterative reconvolution method was able to accurately recover the underlying dynamics of the simulated decays; however, the success of this approach relied greatly on the prior knowledge of the simulations. On the face of it, the four-exponential decay appeared to be adequately described by only three-exponentials; only

**Figure 8.** Residuals associated with (a) the 3-exponential fit and (b) the 4-exponential fit of the 2AP-I fluorescence decay. Upper plots show residuals for all fit channels (~30 ns) while the lower plot shows data for only the first 100 points (~1 ns) of the fit. The red line in the lower plots is a moving average fit of the residuals based on 5 consecutive points. The $\chi_R^2$, Durbin-Watson ($DW$), and Runs Test statistic ($Z$) for each fit is also given.




careful scrutiny of the residuals showed that this fit was actually deficient. This is precisely the reason why the probabilistic methods investigated here may be of significant use in the analysis of time-resolved fluorescence data.

### *4.2. Smoothed exponential series method*

Given the simplicity of the constraint function used, the results obtained from the SESM analysis are quite remarkable. Without imposing any model form on the fit, the SESM analysis gave lifetime distributions that were in very good agreement with the simulated parameters. Of course, there were some limitations to the method, such as the fact that the discrete exponential components were described by peaks rather than points. Additionally, although the SESM method correctly predicted two components for the Exp2G decay, the relative weighting of the components had considerable error. This is hardly surprising though; in addition to the influence of noise, the lifetime components were close in value and the corresponding amplitudes were also fairly similar. The Exp2G system was, in effect, not well suited to probabilistic analysis. In this situation iterative reconvolution is preferred and, indeed, was found to give more accurate fit parameters (though it still suffered from the close proximity of the lifetimes as well as their similar amplitudes, which resulted in a slight discrepancy between the simulated and fitted parameters; a higher signal-to-noise ratio would be required to recover the true decay model).

It is worth pointing out that the optimal value for the regularisation parameter, $\gamma$, for the SESM (and the MEM and BPDN) analysis was estimated through the use of simulated decays, which allowed direct comparison of the fits with the known form of the true decay. In future studies it might be necessary to simulate decays with similar properties to the real data to be fitted (for example, number of channels and signal-to-noise ratio) to provide an estimate for the appropriate regularisation parameter weighting to use. Alternatively, it may be beneficial to use more objective methods, such as the L-curve, minimal product method, or generalized cross-validation techniques, to set the regularisation parameter [45].

### *4.3. Maximum entropy method*

The MEM performed in a comparable manner to the SESM. In general, but not always, the MEM provided narrower peaks for discrete components. It is worth pointing out that, despite longer optimisation times, the MEM has a considerable advantage over the SESM (and other distribution-based analysis strategies) by providing uncorrelated solutions [15]; that is, a-factors within the distribution are not influenced by their neighbours. For example, the peak width obtained by the SESM is influenced by the density of time constants used in the analysis (which changes across the distribution when using logarithmic spacing). In contrast, the solution obtained using the MEM is unaffected by the spacing of time constants. It should be noted that the fitting algorithm used for the MEM analysis was fairly basic. For instance, it did not use strategies such as setting the regularisation parameter with Bayesian logic [46] and stopping the minimisation when the gradients of change in $\chi_R^2$ and entropy ($C_{MEM}$) are orthogonal (since one is minimised while the other is maximised) [17]. Despite these limitations the consistency of the results with other methods suggests that the algorithm performed adequately for the purposes of this study.

### *4.4. Basis pursuit denoising*

The BPDN method appears to be an extremely proficient method for recovering a good estimate of the underlying decay model without imposing any restriction upon the form of the solution. A particularly attractive feature of the BPDN approach is the fact that it attempts to find the solution with fewest components: the simplest model. Again, there are some limits to this method that should be highlighted. As with iterative reconvolution, adequately describing the observed behaviour with a set of discrete lifetimes does not necessarily mean that this is the true model of the system. Indeed, when performing fits of decays that are known to be due to a distribution of lifetimes (for example, Gamma1) it is still possible for a discrete set of lifetimes to successfully model the data

within the resolution and noise limits offered by the experimental technique. As it aims to minimize the number of contributing components, the BPDN analysis approach is inherently poor at recovering a distribution of lifetimes; however, it is encouraging that the general shape of the lifetime amplitudes was consistent with the simulated decay lifetime distribution for the Gamma1 simulation. Additionally, Groma *et al.* found that it was possible to recover a distribution of decay lifetimes using BPDN if a large number of decays (~100) were simulated [3]. As BPDN is sensitive to noise, each decay fit resulted in slightly different peak positions which, on average, recovered the underlying distribution of lifetimes. This is a viable process for simulations, where many decays can be generated instantaneously, but it is unlikely to be feasible for a real system, due to the considerable time it can take to measure even a single decay. More of a concern, however, is the fact that the specific noise present in the decay has influence on the lifetime parameters obtained and, occasionally, the appearance of extraneous components. Similar to the other probabilistic methods, it would seem preferable to refine the solution obtained by BPDN by using a deterministic approach.

## 5. Conclusions

Various analytical methods for time-resolved fluorescence data have been investigated with experimental and simulated decays in an attempt to obtain an overview of their benefits and limitations. The results have provided insight into the relative merits of using deterministic approaches, such as the commonly used iterative reconvolution method, and probabilistic approaches, such as the smoothed exponential series method (SESM), the maximum entropy method (MEM) and recently proposed basis pursuit denoising (BPDN, compressed sensing) method.

In addition to idealised discrete exponentials models, which are typically used to simulate fluorescence decay curves, the analytical methods were also tested with a decay that was based on a distribution of lifetimes. Such a distributive model can be more physically appropriate for some real systems and therefore allowed a more complete assessment of the capabilities of the various methods to be carried out. For the most part, all of the analytical approaches were able to recover the underlying model that had been simulated. The main outcome of this evaluation is that no single method is preferred universally, and there is likely to be value in using a combination of multiple methods when there is ambiguity in the interpretation of the results. While this strategy may still not be sufficient to provide a definitive model of the system, it should be able to provide a model good enough to enable the experimentalist to achieve a satisfactory, physical interpretation of their data.

Ultimately, regardless of the analytical approach taken, the reliability of the results obtained from any method is heavily dependent on the quality of the data available. Care is needed to ensure that potential experimental shortcomings (such as pile up and decay overlap) do not degrade the accuracy of any resulting analysis [47]. Collecting decay data with extremely good signal-to-noise ratio is undoubtedly necessary to be able to discriminate between complex candidate models.

### *5.1. Outlook*

Unfortunately, despite advances in experimental equipment (higher sensitivity detectors, wavelength-independent response, lower background counts, and so on), the fundamental challenge of obtaining the most physically realistic decay model remains, because there is not a technological solution. The main problem with using the null hypothesis to determine the model which best describes the system is that it cannot definitively provide the correct model; it can only indicate that the tested model might be correct. Indeed, it is important to heed the warnings given by James and Ware concerning the interpretation of TCSPC data when there is an absence of supplementary information to corroborate the model chosen, particularly in instances where decays are collected to an inadequate signal-to-noise ratio [10]. Given the unavoidable nature of this problem, all that can be hoped for is that the model that is finally selected is the best possible choice given the information available. The results presented here have shown that the combination of probabilistic and deterministic analysis methods can enable a much more confident prediction to be made about the true model that underlies the observed fluorescence decay.

While clearly powerful, the uptake of probabilistic analysis methods is potentially limited by technical barriers. Fortunately, new tools, which often exploit the wealth of data available from multi-wavelength techniques (such as transient absorption), are being developed. Slavov *et al.* recently introduced OPTIMUS, which is a modular, MATLAB-based package capable of performing global and distribution analysis [45]. Such convenient platforms should facilitate the adoption of probabilistic analysis methods like those described in this study (as well as other advanced analytical techniques); this will be invaluable in making sense of complex systems that would otherwise be difficult to characterise. Furthermore, advances in well-established analysis strategies ensure that the most is made from collected data. Novel approaches should enable faster, more accurate, and more fruitful analysis to be carried out. For instance, a new maximum entropy method algorithm developed by Esposito *et al.* enabled analysis of larger datasets as well as the accurate characterisation of the heterogeneity within a lifetime distribution [48].

Parallel to the development of probabilistic analysis strategies, there should be continued effort to provide appropriate, non-exponential models that accurately

describe the photophysical phenomena that underpin observed experimental data [49]. As has been outlined above, probabilistic models may provide a good first estimate of the underlying model but a more precise (and perhaps more enlightening) description might only be found by using the most physically realistic, deterministic description of the system.

Time-resolved fluorescence techniques provide crucial insight into the underlying mechanics of a great variety of systems. It is therefore imperative that the best is made of the data collected by developing and refining the analysis strategies that are available.

## 6. Acknowledgements

We gratefully acknowledge PhD studentship funding for DAS and GMcK from the Engineering and Physical Sciences Research Council Doctoral Training Account (University of Edinburgh) and the University of Melbourne. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The research data supporting this publication, including MATLAB scripts used for decay analysis, can be accessed at http://dx.doi.org/10.7488/ds/2093. The authors declare that no conflicting interests exist.

## 7. References

[1] Giurleo J T and Talaga D S 2008 Global fitting without a global model: regularization based on the continuity of the evolution of parameter distributions *The Journal of Chemical Physics* **128** 114114

[2] Landl G, Langthaler T, Englt H W and Kauffmann H F 1991 Distribution of event times in time-resolved fluorescence: The exponential series approach—algorithm, regularization, analysis *Journal of Computational Physics* **95** 1-28

[3] Groma G I, Heiner Z, Makai A and Sarlós F 2012 Estimation of kinetic parameters from time-resolved fluorescence data: A compressed sensing approach *RSC Advances* **2** 11481

[4] McKinnon A E, Szabo A G and Miller D R 1977 The deconvolution of photoluminescence data *The Journal of Physical Chemistry* **81** 1564-70

[5] Knight A E W and Selinger B K 1971 The deconvolution of fluorescence decay curves: A non-method for real data *Spectrochimica Acta Part A: Molecular Spectroscopy* **27** 1223-34

[6] Fogarty A C, Jones A C and Camp P J 2011 Extraction of lifetime distributions from fluorescence decays with application to DNA-base analogues *Physical Chemistry Chemical Physics : PCCP* **13** 3819-30

[7] Birch D J S and Imhof R E 1991 *Topics in Fluorescence Spectroscopy,* ed J R Lakowicz (US: Springer) pp 1-95

[8] Box G E P, Jenkins G M and Reinsel G C 1994 *Time Series Analysis: Forecasting and Control* (Englewood Cliffs, NJ: Prentice-Hall)

[9] O'Connor D V and Phillips D 1984 *Time-Correlated Single Photon Counting* (London: Academic Press)

[10] James D R and Ware W R 1985 A fallacy in the interpretation of fluorescence decay parameters *Chemical Physics Letters* **120** 455-9

[11] Bradley J V 1968 *Distribution-Free Statistical Tests,* (Englewood Cliffs, NJ: Prentice-Hall) pp 271-82

[12] Knutson J R, Beechem J M and Brand L 1983 Simultaneous analysis of multiple fluorescence decay curves: a global approach *Chemical Physics Letters* **102** 501-7

[13] Beechem J M and Brand L 1986 Global analysis of fluorescence decay: applications to some unusual experimental and theoretical studies *Photochemistry and photobiology* **44** 323-9

[14] van Stokkum I H M, Larsen D S and van Grondelle R 2004 Global and target analysis of time-resolved spectra *Biochimica et Biophysica Acta* **1657** 82-104

[15] Livesey A K and Brochon J C 1987 Analyzing the distribution of decay constants in pulse-fluorimetry using the maximum entropy method *Biophysical Journal* **52** 693-706

[16] Siemiarczuk A and Ware W R 1989 Temperature dependence of fluorescence lifetime distributions in 1, 3-di (1-pyrenyl) propane with the maximum entropy method *The Journal of Physical Chemistry* **93** 7609-18

[17] Siemiarczuk A, Wagner B D and Ware W R 1990 Comparison of the maximum entropy and exponential series methods for the recovery of distributions of lifetimes from fluorescence lifetime data *Journal of Physical Chemistry* **94** 1661-6

[18] Brochon J-C 1994 Maximum entropy method of data analysis in time-resolved spectroscopy *Methods in Enzymology* **240** 262-311

[19] Swaminathan R and Periasamy N 1996 Analysis of fluorescence decay by the maximum entropy method: Influence of noise and analysis parameters on the width of the distribution of lifetimes *Proceedings of the Indian Academy of Sciences (Chemical Sciences)* **108** 39-49

[20] Hobson M P and Lasenby A N 1998 The entropic prior for distributions with positive and negative values *Monthly Notices of the Royal Astronomical Society* **298** 905-8

[21] Lórenz-Fonfría V A and Padrós E 2005 Maximum Entropy Deconvolution of Infrared Spectra: Use of a Novel Entropy Expression Without Sign Restriction *Applied Spectroscopy* **59** 474-86

[22] Lórenz-Fonfría V A and Kandori H 2007 Bayesian Maximum Entropy (Two-Dimensional) Lifetime Distribution Reconstruction from Time-Resolved Spectroscopic Data *Applied Spectroscopy* **61** 428-43

[23] Ware W R, Doemeny L J and Nemzek T L 1973 Deconvolution of fluorescence and phosphorescence decay curves. Least-squares method *The Journal of Physical Chemistry* **4561** 2038-48

[24] James D R and Ware W R 1986 Recovery of underlying distributions of lifetimes from fluorescence decay data *Chemical Physics Letters* **126** 7-11

[25] Phillips D L 1962 A Technique for the Numerical Solution of Certain Integral Equations of the First Kind *Journal of the ACM (JACM)* **9** 84-97

[26] Liu Y S and Ware W R 1993 Photophysics of polycyclic aromatic hydrocarbons adsorbed on silica gel surfaces. 1. Fluorescence Lifetime Distribution Analysis: An Ill-Condition Problem *The Journal of Physical Chemistry* **97** 5980-86

[27] Cullum J 1979 The effective choice of the smoothing norm in regularization *Mathematics of Computation* **33** 149-70

[28] Gill P R, Wang A and Molnar A 2011 The in-crowd algorithm for fast basis pursuit denoising *Transactions on Signal Processing, IEEE* **59** 4595-605
[29] Chen S S, Donoho D L and Saunders M A 1998 Atomic Decomposition by Basis Pursuit *SIAM Journal on Scientific Computing* **20** 33-61
[30] Candès E J, Romberg J K and Tao T 2006 Stable signal recovery from incomplete and inaccurate measurements *Communications on Pure and Applied Mathematics* **LIX** 1207-23
[31] Candès E and Wakin M B 2008 An Introduction to Compressive Sampling *Signal Processing Magazine, IEEE* **25** 21-30
[32] Donoho D L 2006 Compressed sensing *IEEE Transactions on Information Theory* **52** 1289-306
[33] Tibshirani R 1996 Regression shrinkage and selection via the lasso *Journal of the Royal Statistical Society Series B-Statistical Methodology* **58** 267-88
[34] Zou H and Hastie T 2005 Regularization and variable selection via the elastic net *Journal of the Royal Statistical Society Series B-Statistical Methodology* **67** 301-20
[35] Damayanti N P, Craig A P and Irudayaraj J 2013 A hybrid FLIM-elastic net platform for label free profiling of breast cancer *The Analyst* **138** 7127-34
[36] Dorlhiac G F, Fare C and van Thor J J 2017 PyLDM - An open source package for lifetime density analysis of time-resolved spectroscopic data *PLOS Computational Biology* **13** e1005528-e
[37] Bajzer Ž, Therneau T M, Sharp J C and Prendergast F G 1991 Maximum likelihood method for the analysis of time-resolved fluorescence decay curves *European biophysics journal : EBJ* **20** 247-62
[38] Enderlein J, Goodwin P M, van Orden A, Ambrose W P, Erdmann R and Keller R A 1997 A maximum likelihood estimator to distinguish single molecules by their fluorescence decays *Chemical Physics Letters* **4** 464-70
[39] Somsen O J G, van Hoek A and van Amerongen H 2005 Fluorescence quenching of 2-aminopurine in dinucleotides *Chemical Physics Letters* **402** 61-5
[40] Menezes F, Fedorov A, Baleizão C, Valeur B and Berberan-Santos M N 2013 Methods for the analysis of complex fluorescence decays: sum of Becquerel functions versus sum of exponentials *Methods and Applications in Fluorescence* **1** 015002
[41] Donoho D, Drori I, Stodden V, Tsaig Y and Shahram M 2007 *Version 2.0, See; http://sparselab.stanford.edu/*
[42] Lackowicz J R 2006 *Principles of Fluorescence Spectroscopy* (US: Springer Science & Business Media, LLC)
[43] Gopich I V and Szabo A 2012 Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET *Proceedings of the National Academy of Sciences of the United States of America* **109** 7747-52
[44] Jones A C and Neely R K 2015 2-aminopurine as a fluorescent probe of DNA conformation and the DNA–enzyme interface *Quarterly Reviews of Biophysics* **48** 244-79
[45] Slavov C, Hartmann H and Wachtveitl J 2015 Implementation and Evaluation of Data Analysis Strategies for Time-Resolved Optical Spectroscopy *Analytical Chemistry* **87** 2328-36
[46] Skilling J and Gull S F 1991 Bayesian maximum entropy image reconstruction *Lecture Notes-Monograph Series* **20** 341-67
[47] Henry E, Deprez E and Brochon J-C 2014 Maximum entropy analysis of data simulations and practical aspects of time-resolved fluorescence measurements in the study of molecular interactions *Journal of Molecular Structure* **1077** 77-86
[48] Esposito R, Altucci C and Velotta R 2013 Analysis of simulated fluorescence intensities decays by a new Maximum Entropy Method algorithm *Journal of Fluorescence* **23** 203-11
[49] Boens N and Van der Auweraer M 2014 Identifiability of models for time-resolved fluorescence with underlying distributions of rate constants *Photochemical & Photobiological Sciences* **13** 422-30